

FIELD OF THE INVENTION

The present invention relates to a method of forming, from data obtained by exploration of a zone of a heterogeneous medium, a model representative of the distribution in the zone of a physical quantity (at least partly) free of the presence of correlated noises that may be contained in the data.

The method applies for example to the quantification of the acoustic impedance in an underground zone.

BACKGROUND OF THE INVENTION

The process consisting in seeking a model that adjusts to experimental measurements has developed in nearly all the fields of the sciences or technology. Such an approach is known under various names : least-squares method for parameters estimation, for inverse problem solution. For a good presentation of this approach within the context of geosciences, one may for example refer to :

- Tarantola, A. : « Inverse Problem Theory : Method for Data Fitting and Model Parameter Estimation », Elsevier, Amsterdam, 1987.

It can be noted that the term « lest squares » refers to to the square of the norm in the data space for quantifying the difference between the response of a model (which is the image of the model by a previously selected modelling operator) and the data, a cost function that has to be minimized to solve the problem. Using the square of the norm to define the cost function is just a practical convenience, and it is not fundamentally essential. Besides, many authors use, for various reasons, another definition of the cost function but this definition remains based on the use of the norm, or of a semi-norm, in

the data space. Finally, we have considerable latitude for selecting the norm (or the semi-norm) in the data space (we are in no way forced to use the Euclidean norm). In the case of noise-containing data, the solution can substantially depend on the choice made at this stage. For more developments concerning this problem, one can for example refer to the following publications :

- Tarantola, A. : « Inverse Problem Theory : Method for Data Fitting and Model Parameter Estimation », Elsevier, Amsterdam, 1987 ; Renard and Lailly, 2001 ; Scales and Gersztenkorn, 1988 ; Al-Chalabu, 1992.

The measuring results often contain errors. Modelling noises add further to these measuring errors when the experimentation is compared with modelling results : modellings are never perfect and therefore always correspond to a simplified view of reality. The noise will therefore be described hereafter as the superposition of :

- non correlated components (white noise for example),
- correlated components, which means that the existence of a noise on a measuring sample translates into the existence of a noise of equal nature on certain neighbouring measuring points ; modelling noises typically belong to this category.

When the data contain correlated noises, the quality of the model estimated by solving the inverse problem can be seriously affected thereby. As already mentioned, no modelling operator is perfect. It is therefore the work of the whole community of the people involved in the identification of parameters describing a model which is hindered by the existence of correlated noises. Among these people, seismic exploration practitioners are among the most concerned ones : in fact, their data have a poor or even very bad signal-to-noise ratio. This is the reason why correlated noise filtering

techniques are an important part of seismic data processing softwares. The most conventional techniques use a transform (Fourier transform for example) where signal and noise are located in different areas of space, thus allowing separation of the signal and of the noise. For a general presentation of the conventional methods for noise
 5 elimination on seismic data, one may refer to the book by Yilmaz :

- Yilmaz, O. 1987 : « Seismic data processing », Investigation in Geophysics No.2, Society of Exploration Geophysicists, Tulsa, 1987.

However, these techniques are not perfect : they presuppose that a transformation allowing complete separation of the signal and of the noise has been found. In this
 10 context, correlated noises are particularly bothersome because they can be difficult to separate from the signal (which is also correlated) and can have high amplitudes. It is therefore often difficult to manage the following compromise : either the signal is preserved, but a large noise residue remains, or the noise is eliminated, but then the signal is distorted. These filtering techniques can be implemented before solving the
 15 inverse problem : they constitute then a preprocessing applied to the data. The quality of the inverse problem solution then widely depends on the ability of the filters to eliminate the noise without distorting the signal.

The approach introduced by Nemeth et al. in the following publication :

- Nemeth T. et al. (1999), « Least-Square Migration of Incomplete Reflection Data »,
 20 Geophysics, 64, 208-221

constitutes an important advance for the inversion of data disturbed by high-amplitude correlated noises. These authors propose eliminating the noise by solving an inverse problem : after defining the space of the correlated noises as the image space of a vector

space B (space of the noise-generating functions) by a linear operator T and the signal as the image space of a vector space M (models space) by a modelling operator F (assumed to be linear), they seek the signal in $F(M)$ and the noise in $T(B)$ whose sum is as close as possible (in the sense of the Euclidean norm or, on a continuous basis, of the norm L^2) to the measured datum. This technique constitutes an important advance insofar as it allows elimination (or at least very substantial reduction) of high-amplitude correlated noises (surface waves for example) that are difficult to separate from the signal by means of conventional filtering techniques. However, according to these authors, an increase by an order of magnitude of the computing time required for solution of the conventional inverse problem, i.e. without seeking the correlated component of the noise, is the price to pay for these performances. Furthermore, the result obtained with the method is extremely sensitive to any inaccuracy introduced upon definition of operator T : this is the ineluctable compensation for the high aptitude of the method to discriminate signal and noise.

SUMMARY OF THE INVENTION

The method according to the invention allows to estimate, from data obtained by exploration of an underground zone, a model representative of the distribution, in the zone, of at least one physical quantity, this model being amply free of the presence of correlated noises that may be contained in the data. The method essentially comprises the following stages :

a) acquisition of measurements giving information about certain physical characteristics of the zone by following a predetermined experimental protocol,

b) specification of a modelling operator which associates, with a model of each physical quantity, synthetic data that constitute the response of the model, the measurements and the synthetic data belonging to a data space,

c) for each correlated noise referenced by a subscript j ranging from 1 to J ,
 5 selection of a modelling operator which associates a correlated noise with a noise-generating function belonging to a predetermined space of noise-generating functions,

d) specification of a norm or of a semi-norm in the data space,

e) specification of a semi-norm in the noise-generating functions space for which each noise modelling operator establishes substantially an isometric relation between
 10 the noise-generating functions space and the data space,

f) definition of a cost function quantifying the difference between the measurements on the one hand and the superposition of the model response and of the correlated noises associated with the noise-generating functions on the other hand, and

g) adjustment of the model and of the noise-generating functions by minimizing the
 15 cost function, by means of an algorithmic method taking advantage of the isometry properties of the noise modelling operators.

According to a first implementation mode, the distribution as a function of the depth of the acoustic impedance in the medium is sought, the correlated noises affecting the data are tube waves identified each by parameters characterizing their propagation, the
 20 measured data are VSP type data obtained by means of pickups suited to detect the displacement of particles in the medium in response to a localized seismic excitation, the location of the pickups, the recording time and the time sampling points being defined, and the modelling operator selected associates the synthetic data with an

acoustic impedance distribution as a function of the evaluated depth in traveltime and with the vertical stress measured as a function of time at the depth of the first pickup.

The cost function selected to quantify the difference between the measurements on the one hand and the superposition of the model response and of the correlated noises associated with the noise-generating functions on the other hand is for example the square of the semi-norm of this difference in the data space.

According to an embodiment, adjustment of the model and of the noise-generating functions is for example obtained by means of a relaxation method for eliminating the unknowns corresponding to each correlated noise generating function, this relaxation method being implemented within the iterations of a quasi-Newtonian algorithm for calculation of the model.

According to an embodiment, numerical calculation of the image of a model by this operator is carried out for example by numerical solution of the 1D waves equation for the model considered, by selecting values taken by the displacement of the particles at the locations of pickups and at the previously defined time sampling points, and by applying an operator likely to compensate for the spherical divergence and attenuation effects.

According to an embodiment, the numerical noise modelling operator is a finite-difference centered numerical scheme for discretizing the noise transport equation, and the noise-generating function involved as the initial condition along the edge of said zone associates with each correlated noise a noise-generating functions space consisting of the support time functions in a given time interval.

Various examples of norms or semi-norms selected for the data space and the noise-generating functions space are given in the detailed description hereunder.

According to a second implementation mode, the distribution of disturbances, in relation to a previously selected reference medium, of the impedance and of the velocity in said zone of the medium is sought, the correlated noises affecting the data are due to multiple reflections whose kinematics and amplitude variations with the offset have been previously estimated, the measured data are picked up by seismic surface pickups, the location of said pickups, the seismic excitation mode, the recording time and the time sampling points being defined, and the modelling operator being defined via a linearization of the waves equation around the reference medium.

The cost function selected to quantify the difference between the measurements on the one hand and the superposition of the model response and of the correlated noises associated with the noise-generating functions on the other hand is for example the square of the semi-norm of this difference in the data space.

According to an embodiment, adjustment of the model and of the noise-generating functions is obtained by means of a block relaxation method for eliminating the unknowns corresponding to each correlated noise generating function, this relaxation method being implemented within the iterations of a conjugate gradient algorithm for calculation of the model.

BRIEF DESCRIPTION OF THE FIGURES

Other features and advantages of the method according to the invention will be clear from reading the description hereafter of an embodiment given by way of non limitative example, with reference to the accompanying drawings wherein :

- 5 - Figure 1 shows an example of data obtained by means of a vertical seismic prospecting method (VSP),
- Figure 2 shows a model of distribution of the acoustic impedance with depth,
- Figure 3 shows synthetic VSP data obtained on the basis of the impedance model of Fig.2,
- 10 - Figure 4 shows an example of VSP data contaminated by a single correlated noise,
- Figure 5 shows an example of VSP data contaminated by two correlated noises,
- Figure 6 shows the impedance distribution as a function of depth, obtained by inversion of the noise-containing data of Fig.4,
- Figure 7 shows the inversion residues (differences between the data of Fig.4 and the
15 seismic response of the model of Fig.6),
- Figure 8 shows the impedance distribution as a function of depth, obtained by inversion of the noise-containing data of Fig.5,
- Figure 9 shows the corresponding inversion residues (differences between the data of Fig.5 and the seismic response of the model of Fig.8),
- 20 - Figure 10 shows the impedance distribution as a function of depth, obtained by inversion of the noise-containing data of Fig.4 by seeking the correlated noise in form

of the superposition of two correlated noises having inaccurate propagation properties, i.e. different from those of the noise appearing in the data of Fig.4,

- Figure 11 shows the corresponding inversion residues (differences between the data of Fig.4 and the seismic response of the model of Fig.10),
- 5 - Figures 12A, 12B respectively show seismic data with multiple reflections and the seismic response of the model obtained after conventional linearized inversion,
- Figures 13A, 13B respectively show an example of impedance distribution and of propagation velocity for which the seismic data of Fig.12A constitute the seismic response, and
- 10 - Figures 14A, 14B respectively show the comparison of the seismic responses of the models obtained by conventional inversion (14A) and by inversion according to the proposed method involving picking of the moveout of the multiple and an estimation of the amplitude variations with the offset (14B).

DETAILED DESCRIPTION

15 Formulation of the problem

We have data resulting from a sampled measurement of a function. This function depends on several variables (space or space-time variables for example).

These data, called d , correspond to measurements carried out to obtain information about a model m . They contain various noises :

- 20 - correlated noises (or a superposition of correlated noises),
- non correlated noises.

The problem is to determine quantitatively model m (or functions of this model) from datum d .

We therefore select a modelling operator F (linear or not) which associates with model m the response of the model. This operator actually models the real physical phenomenon only imperfectly. This is essentially (but not exclusively) the reason why correlated noises appear in the data : these noises correspond to a signal related to the model but the relation between them appears to be too complex to be included in modelling operator F , which we want to keep relatively simple for various reasons.

Under such conditions, we model the noise by introducing one or more noise-generating spaces B_j (j ranging from 1 to J) and the associated noise modelling operators called T_j . In order to find model m from data d , we propose seeking this model as the solution to the optimization problem.

$$\min_{(m, \beta_1, \beta_2, \dots, \beta_J) \in M \times \prod_{j=1}^J B_j} C(m, \beta_1, \beta_2, \dots, \beta_J) = \left\| F(m) + \sum_{j=1}^J T_j \beta_j - d \right\|_D^2 \quad (1)$$

where $\| \cdot \|_D$ is a norm (or possibly a semi-norm) in the data space.

The above formalism differs from the approach proposed by Nemeth et al. (1999) in that :

- modelling operator F is not necessarily linear,
- we are justified in considering a superposition of correlated noises of different types (in the sense that they are modelled by different operators referenced by subscripts j), at the price of a marked increase in the number of calculations concerning the number of

unknowns (which particularly complicates the solution of the optimization problem) as well as the number of noise modellings to be carried out,

- we reserve the possibility of selecting norm $\| \cdot \|_D$ or possibly the semi-norm to suit our convenience in the data space.

5 The choice that we are going to make for $\| \cdot \|_D$ will be imposed by considerations linked with the efficiency of the solution method (described below), this efficiency allowing to deal with the case of a correlated noise superposition. This is an important opening : besides the possibility of processing data containing correlated noises with complex characteristics, we can in this way overcome the difficulties encountered in
10 Nemeth et al.'s approach, linked with the high sensitivity of the result to an inaccuracy introduced upon definition of the noise modelling operator : to overcome this difficulty, we just have to seek the noise by superposition of noises associated with approximate modelling operators T_j .

It can be noted that selection of the square of norm $\| \cdot \|_D$ involved in the
15 optimization problem is not essential : we do not change the solution by selecting instead a function which, like the square function, is an increasing function on \Re^+ (or a decreasing function provided that min is changed into max).

The solution method

The method consists in suitably selecting :

- 20 - norm (or semi-norm) $\| \cdot \|_D$ in the data space,
- norms $\| \cdot \|_D$ in each space B_j

so as to obtain a high-performance algorithmic solution to optimization problem (1).

The expression « suitably select » means see to it that each operator T_j constitutes an isometry (or an isometry approximation) for norms $\| \cdot \|_j$ and $\| \cdot \|_D$ respectively in the initial space and in the end space. The main idea in the determination of norms allowing
 5 to make each operator T_j isometric is based on the existence of energy conservation type equalities verified by the solutions to the partial differential equations (or of discrete energy equalities for the solutions to certain numerical schemes used for discretization of these equations).

If we can make such choices, it is possible to make the cost function minimization
 10 very simple by eliminating the unknowns associated with the correlated noise generating functions (determination of Schur's complement). This elimination is performed by means of a suitable algorithm such as the block relaxation method, a block being associated with a noise-generating function, or the conjugate gradient method with symmetrical-block Gauss-Seidel preconditioning, etc., all of them algorithmic
 15 implementation methods well-known to the man skilled in the art, presented for example by :

- Golub, G.H. et al. (1983) « Résolution numérique des grands systèmes linéaires » (Eyrolles).

Simplified minimization of the cost function allows to considerably reduce the time
 20 required for numerical solution.

First implementation example applied to 1D inversion of VSP data contaminated by a tube wave

VSP (Vertical Seismic Profile) data belong to conventional well survey acquisitions. They have multiple applications, one of the main ones being their use for
 5 establishing a connection between seismic surface data and logging measurements. The inversion of VSP data as described for example by :

- Macé, D. and Lailly, P. (1984) A Solution of an Inverse problem with 1D Wave Equation applied to the Inversion of Vertical Seismic Profile ; Proc. of the 6th Intern. Conf. In « Analysis and Optimisation of Systems », Nice, 2, 309-323

10 has been developed to that end. In the 1D inverse problem, we assume that the subsurface model is laterally invariant and that a plane-wave excitation propagates vertically. The unknowns of the problem are the acoustic impedance distribution as a function of depth (measured in vertical traveltimes and over an interval starting at the depth of the first receiver) and the seismic excitation mode (precisely the time function
 15 characterizing the boundary condition at the depth of the first pickup). This is a non linear inverse problem, the VSP data depending non linearly on the impedance distribution.

VSP data are often contaminated by a correlated noise of very high amplitude : the fluid wave. This wave is guided in the mud that invades the well. Its propagation
 20 velocity is low in relation to the propagation velocity in the rocks surrounding the well. Furthermore, the propagation velocity of this wave depends on the frequency (dispersive propagation). Finally, this wave reflects notably at the well bottom and it furthermore gives rise to another propagation mode. Typical VSP data are given in the

figure below : the main characteristics of the tube wave can be seen. Its amplitude and the extension of the contaminated zone make it difficult to use the signal contained in the many seismic samples contaminated by the fluid wave.

We describe hereafter the implementation of the method for 1D inversion of the VSP data.

Experimental protocol

It consists in specifying :

- the various depths at which the receivers are positioned in the well (assumed to be vertical) : in our experiments, the pickups cover the range of depths measured in one-way time $[x_{\min} = 0.05\text{s} ; x_{\max} = 0.2\text{s}]$, with a receiver every $\Delta x = 1 \text{ ms}$ (in one-way time) ; we thus obtain samples numbered from 0 to I,
- the physical nature of the measurement performed by these pickups : in our experiments, the pickups measure the vertical displacement resulting from the wave propagation,
- the recording time : in our experiments, the pickups measure the vibrational state over the time interval $[0, T = 1.5\text{s}]$, these data being sampled every $\Delta t = 1 \text{ ms}$. We thus obtain samples numbered from 0 to N.

We call d_i^n the data sample recorded at the depth $x_{\min} + i \Delta x$ and at the time $n \Delta t$. We have of course : $T = N \Delta t$ and $x_{\max} = x_{\min} + I \Delta x$.

Acquisition of data in accordance with this experimental protocol

Our experiments were carried out from synthetic data, i.e. obtained by numerical modelling, itself carried out in two stages :

Creation of synthetic data containing no noise : this creation was performed by numerical solution of the 1D acoustic wave equation, the model being the acoustic impedance distribution shown in Fig.2 (the depth is measured in one-way time) and the excitation mode (Neumann's boundary condition at the depth 0) being a conventional
 5 Ricker wavelet.

The synthetic VSP data resulting from this modelling are shown in Figure 3 : the vertical axis representing the observation time is graduated from 0 to 1.5 s and the horizontal axis representing the depth (in one-way travelttime) of the various pickups is graduated from 0.05 to 0.2 s.

10 *Addition of noise to the data*

We are going to disturb the VSP data thus calculated by adding thereto, on the one hand, a random noise (but filtered in the seismic frequency band) of rather high amplitude and, on the other hand, one or more correlated noises of very high amplitude. Each noise will propagate downwards with a low velocity which depends on the
 15 frequency (dispersive propagation) : these correlated noises are supposed to represent fluid waves. They have been modelled by numerical solution (use of the finite-difference centered scheme) of a transport equation with a constant propagation velocity. Dispersion of the waves has been obtained using large discretization intervals in the finite-difference scheme. In the case of the superposition of several correlated
 20 noises, the propagation velocities associated with each noise are different. Figures 4, 5 show the VSP data contaminated by a single correlated noise (Fig.4) and by the superposition of two correlated noises (Fig.5).

Specification of the modelling operator

Modelling is carried out by solving the wave equation :

$$\sigma(x) \frac{\partial^2 y}{\partial t^2} - \frac{\partial}{\partial x} \left(\sigma(x) \frac{\partial y}{\partial x} \right) = 0 \text{ in }]x_{\min}, +\infty[\times]0, T]$$

with Neumann's boundary condition :

$$-\sigma(x_{\min}) \frac{\partial y}{\partial x}(x_{\min}, t) = h(t) \text{ on }]0, T]$$

and the initial conditions :

$$y(x, 0) = \frac{\partial y}{\partial t}(x, 0) = 0 \text{ on } [x_{\min}, +\infty]$$

5

Definition of the modelling operator first requires definition of the observation operator which formalizes the experimental protocol described in §1. The observation operator thus is the operator which associates with function $y(x, t)$, solution to the above system, samples $y(x_i, t^n)$ for $x_i = x_{\min} + i \Delta x$ and $t^n = n \Delta t$, $i=0, \dots, I$; $n=0, \dots, N$.

10 The modelling operator that we call $F(m)$ thus is the (non linear) operator which associates with function pair $m=(\sigma(x), h(t))$ samples $y(x_i, t^n)$ for $x_i = x_{\min} + i \Delta x$ and $t^n = n \Delta t$, $i=0, \dots, I$; $n=0, \dots, N$.

Specification of the modelling operator associated with each correlated noise

15 We specify in this paragraph the procedure that can be used for modelling each correlated noise. Each correlated noise will be characterized by correlation directions specific thereto, which we assume to be known, at least approximately : these correlation directions are specified by means of a field of correlation vectors $\vec{c}_j(x, t)$ of components $(c_j^x(x, t), c_j^t(x, t))$ which has to be defined at any point of the domain $[x_{\min},$

$x_{\max}]X[0,T]$. We can, in an equivalent way, specify the correlation directions by means of a pencil of correlation lines that will represent the field lines associated with $\vec{c}_j(x,t)$. Such correlation lines can be obtained according to the technique described in patents EP-354,112 (US-4,972,383) filed by the applicant, from a pick of several phases characteristic of the noise, the information being extended to the whole domain $[x_{\min}, x_{\max}]X[0,T]$ by means of conventional interpolation and/or extrapolation procedures. Specification of the correlation directions amounts to specifying the propagation velocity distribution of the noise $c_j(x,t)$ that is related to the correlation vector by the relation $c_j(x,t) = \frac{c_j^x(x,t)}{c_j^t(x,t)}$. We assume here that there are no wave amplitude variations during propagation of the wave (a different situation is presented for the 2nd application example). Modelling of a correlated noise as specified above is based on the following observation : a wave propagating along correlation lines without amplitude change is solution to the transport equation :

$$\vec{\nabla} b(x,t) \cdot \vec{c}(x,t) = 0$$

15 In this context, we introduce :

- a space B_j of noise-generating functions which will be the space of support functions $\beta_j(t)$ in $[t_j^{\min}, t_j^{\max}] \subset [0,T]$ and that we extend by 0 in the whole interval $[0,T]$,

- noise modelling operator T_j :

$$T_j : \beta_j(t) \in B_j \rightarrow b_j(x,t) \in D$$

20 solution to the transport equation

$$\bar{\nabla} b_j(x, t) \cdot \bar{c}_j(x, t) = 0 \text{ in } [x_{\min}, x_{\max}] \times [0, T]$$

and meeting the initial conditions :

$$\begin{cases} b_j(x, t = 0) = 0 \\ b(x = x_{\min}, t) = \beta_j(t) \end{cases}$$

It can be noted that the geometry of the pencil of correlation lines (or, which
 5 amounts to the same thing, the propagation velocity distribution) cannot be any
 geometry : the correlation lines cannot intersect, or the transport equation solution
 problem would not be correctly set. For the same reasons, the correlation lines cannot
 intercept twice the edges on which the initial conditions are specified. The latter
 consideration can lead (in order to model a correlated noise corresponding to an
 10 upgoing wave) to specify the initial condition no longer at $x=x_{\min}$ but at $x=x_{\max}$ or even at
 an intermediate value, even if it means decomposing Cauchy's problem into two
 problems associated with the domains located on either side of this intermediate edge.

We furthermore assume that the geometry of the correlation lines is such that (more
 general situations can be considered, but the following results then take a different
 15 form) :

$$c_j(x, t) = \xi_j(x) \times \tau(t)$$

$$\text{ou encore : } c_j^x(x, t) = \xi_j(x) \text{ et } c_j^t(x, t) = \frac{1}{\tau(t)}$$

Finally, we assume that the pencil of correlation lines that intercept at $x=x_{\min}$
 interval $[t_j^{\min}, t_j^{\max}]$ do not intercept at $t=T$ interval $[x_{\min}, x_{\max}]$.

In fact, the transport equation is solved by means of a numerical scheme. It is for example possible to use the conventional finite-difference centered scheme. We therefore select discretization intervals $\Delta x_j'$ and $\Delta t_j'$ (which are not necessarily equal to Δx and Δt , but which we assume to be submultiples of these quantities, even if more general situations can be considered) and we introduce a grid whose nodes are the points of coordinates $(x_{\min} + i'\Delta x_j', n'\Delta t_j')$ with i' and $n' \in \mathbb{N}$. The conventional finite-difference centered scheme explained hereunder allows, from the initial conditions, to calculate step by step the different values taken by function $b_j(x,t)$ at the various nodes of the grid (to simplify the notations, we have left out subscript j associated with the correlated noise considered).

$$\frac{\xi_{i'+\frac{1}{2}}}{2\Delta t'} \left[b_{i'+1}^{n'+1} + b_{i'}^{n'+1} - b_{i'+1}^{n'} - b_{i'}^{n'} \right] + \frac{1}{2\Delta x' \tau^{n'+\frac{1}{2}}} \left[b_{i'+1}^{n'+1} + b_{i'+1}^{n'} - b_{i'}^{n'+1} - b_{i'}^{n'} \right] = 0$$

In the above formula, quantity $\alpha_{i',n'}$ represents the evaluation of quantity α (α is a generic term) at the point of coordinates $(x_{\min} + i'\Delta x_j', n'\Delta t_j')$.

Of course, the use of discretization intervals $\Delta x_j'$ and $\Delta t_j'$ sufficiently small in relation to the wavelength is required if we want the numerical scheme to give a precise approximation of the function solution to the transport equation.

However, the use of larger discretization intervals (but still smaller than the wavelength) allows to model more complex propagation phenomena by making the wave propagation velocity dependent on the frequency : dispersive propagations such as

the propagation of tube waves can thus be modelled. Selection of suitable discretization intervals for modelling a particular mode of the tube wave can be made :

By determining, after examining the data, the propagation velocity of this mode (which depends here not only on space and time, but also on the frequency) ;

5 tomography techniques can be used to that end, such as those presented in the following publication :

- Ernst, F. et al. (2000) ; Tomography of Dispersive Media ; J. Acoustic Soc. Am., 108, 105-116.

By knowing the dispersive propagation properties associated with the numerical
10 scheme, properties that follow from an analysis of the numerical dispersion relation, as described for example in the following publication :

- Alford, R.M. et al. (1974), Accuracy of Finite Difference Modeling of the Acoustic Wave Equation ; Geophysics, 6, 834-842.

Specification of a norm or of a semi-norm in the data space

15 A first choice consists in taking as the norm in the data space any discrete norm which is meant to be an approximation (via a quadrature formula) to the norm defined on a continuous basis by :

$$\|u\|_{\mathcal{D}} = \left(\int_{x_{min}}^{x_{max}} dx \int_0^T dt \frac{1}{\tau(t)} u^2(x, t) \right)^{\frac{1}{2}}$$

We can also use the semi-norm, which is a better choice as we shall see in the next paragraph :

$$\|u\|_{\mathcal{D}} = \left(\Delta x \Delta t \sum_{i=0}^{I-1} \sum_{n=0}^{N-1} \frac{1}{\tau^{n+\frac{1}{2}}} (u_i^{n+1} + u_i^n)^2 \right)^{\frac{1}{2}}$$

where u is any vector of the data space (i and n are the indices representing the sample numbers respectively in space and in time).

Besides, other choices are also possible.

Specification, in each space B_j of noise-generating functions, of a norm for which each noise modelling operator T_j establishes an isometric relation (or the approximation of an isometric relation) between the noise-generating functions space and the data space (provided with the semi-norm defined in §5)

We specify in this paragraph the procedure that can be used to define the norm in the noise-generating functions space associated with each correlated noise. This procedure being the same for each noise, we describe it generically and leave out, in the formulas, subscript j associated with the noise considered.

A first choice consists in providing each noise-generating functions space B with any discrete norm meant to be an approximation (via a quadrature formula) to the norm defined on a continuous basis by :

$$\|\beta\|_B = \left(X \int_0^T dt \frac{1}{\tau(t)} \beta^2(t) \right)^{\frac{1}{2}}$$

in which case, using the fact that, by means of the hypotheses set out in §4, we have $b_j(x,T)=0$ for $x \in [x_{\min}, x_{\max}]$, linear operator T_j performs an isometry between B_j and D . However, since calculation of $T_j \beta_j$ is carried out numerically, the isometric character of operator T will not be perfect but only approximate, the approximation being all the better as discretization intervals Δx , Δt , $\Delta x'_j$, $\Delta t'_j$ are small. The fact that we then have only approximately an isometry will lead to lower performances concerning the algorithmic implementation presented in paragraph 8 hereafter.

A better choice consists in providing each noise-generating functions space B_j with the semi-norm :

$$\|\beta\|_B = \left(\Delta x \Delta t I \sum_{n=0}^{N-1} \frac{1}{\tau^{n+\frac{1}{2}}} (\beta^{n+1} + \beta^n)^2 \right)^{\frac{1}{2}}$$

10

In this case, operator T_j rigorously performs an isometry between B_j and D , at least if $u_i^N = 0 \forall i = 0, I-1$. This results from the discrete energy equality :

$$\begin{aligned} I \sum_{n=0}^{N-1} \frac{1}{\tau^{n+1/2}} [u_0^{n+1} + u_0^n]^2 &= \sum_{i=0}^{I-1} \sum_{n=0}^{N-1} \frac{1}{\tau^{n+1/2}} [(u_i^{n+1} + u_i^n)^2] \\ &+ \sum_{i=0}^{I-1} \sum_{k=0}^i \frac{\Delta x \xi_{k+1/2}}{\Delta t} [(u_{k+1}^N + u_k^N)^2] - \frac{I \Delta x (u_0^N)^2 \xi_{-1/2}}{\Delta t} \end{aligned}$$

an equality valid in the case (simplified for presentation's sake) where $\Delta x' = \Delta x$ and $\Delta t' = \Delta t$ and if $\xi_{-1/2} = \xi_{1/2}$ has been defined.

15

Definition of the cost function

We define the cost function by the formula as follows :

$$C(m, \beta_1, \beta_2, \dots, \beta_J) = \left\| F(m) + \sum_{j=1}^J T_j \beta_j - d \right\|_D^2$$

Seeking the model and the noise-generating functions minimizing the cost function, this search being carried out by an algorithmic method taking advantage, explicitly or implicitly, of the isometry properties of the noise modelling operators (see 6)

- 5 The algorithmic method uses the specificity of Schur's complement, a specificity associated with the isometric character of operators T_j . This can be done for example by realizing that minimizing $C(m, \beta_1, \beta_2, \dots, \beta_J)$ amounts to minimizing $C'(m) = C(m, \tilde{\beta}_1(m), \tilde{\beta}_2(m), \dots, \tilde{\beta}_J(m))$ where $(\tilde{\beta}_1(m), \tilde{\beta}_2(m), \dots, \tilde{\beta}_J(m))$ minimizes $C(m, \beta_1, \beta_2, \dots, \beta_J)$ for given m . It can be noted that the Hessian associated with this
- 10 quadratic form is Schur's complement. $C'(m)$ can be minimized by means of any optimization method such as the BFGS version of a quasi-Newtonian method.

- Determination of $(\tilde{\beta}_1(m), \tilde{\beta}_2(m), \dots, \tilde{\beta}_J(m))$ is very easy thanks to the isometric character of operators T_j , notably if choices (2) and (3) have been respectively selected to define $\| \cdot \|_D$ and $\| \cdot \|_{B_j}$. If there is only one type of noise ($J=1$) and if the isometry is
- 15 perfect, determination of $\tilde{\beta}_1(m)$ simply requires evaluation of the gradient of the quadratic form. If there is a superposition of several noises ($J>1$), various algorithms can be considered to take advantage of the isometric character of operators T_j . $(\tilde{\beta}_1(m), \tilde{\beta}_2(m), \dots, \tilde{\beta}_J(m))$ can for example be determined by means of a block relaxation method (Gauss-Seidel alias), a block being associated with a group of unknowns
- 20 $\tilde{\beta}_j(m)$, a relaxation iteration simply requiring (still in the case of a perfect isometry) a single evaluation of the gradient of the quadratic form associated with the block

considered. For problems made difficult by the proximity of the correlation directions of the various noise types, a preconditioned conjugate gradient method can be considered, the preconditioning matrix being of symmetrical block (relaxation alias) Gauss-Seidel type. Then again, preconditioning is very easily implemented, solution of the problem associated with a block only requiring evaluation of the gradient of the quadratic form associated with the block considered.

Figures 6 to 11 show the results obtained with three types of experiments using the method.

In the case of an inversion of noise-containing data (Fig.4) both by superposition of a correlated noise and of a random noise, Fig.6 gives the impedance distribution obtained, and Fig.7 shows the inversion residues.

In the case of an inversion of noise-containing data (Fig.5) both by superposition of two correlated noises and of a random noise, Fig.8 gives the impedance distribution obtained, and Fig.9 shows the inversion residues.

In the case of an inversion of noise-containing data (Fig.4) both by superposition of a correlated noise and of a random noise where the propagation properties of the fluid wave are not precisely known, it is natural to seek a correlated noise model consisting of the superposition of two correlated noises having inaccurate propagation velocities but by which the true propagation velocity of the fluid wave is framed. Fig.10 gives the impedance distribution obtained, and Fig.11 shows the inversion residues.

Second example of application of the method to determination of a model by linearized inversion of multi-offset data contaminated by multiple reflections

This application differs from the first one mainly in that it illustrates the possibilities afforded by the method for taking into account noise amplitude variations along the correlation directions. Another difference lies in the propagation properties of the correlated noises, propagation of the multiple reflections being not dispersive.

5 Linearized inversion is one of the sophisticated methods used by geophysicists to obtain a quantitative model of subsurface heterogeneities, this type of information being precious notably for reservoir characterization. The data given here are surface seismic data. This method is described for example in the following publication :

- Bourgeois, A. et al. (1989) « The Linearized Seismic Inverse Problem : an
- 10 Attractive Method for a Sharp Description of Strategic Traps » : 59th Ann. Intern. Mtg. Soc. Expl. Geoph. Expanded Abstract, pp.973-976.

Furthermore, it is necessary to have a reference model, consisting (in the context of an acoustic model) of a velocity distribution and of an acoustic impedance distribution : it is the model around which a linearization will be carried out in the waves equation

15 (Born's approximation). In order to simplify the presentation, we take the 1D case but this hypothesis is all the less essential as the practical applications mainly concern 3D seismic methods. In the 1D context, all of the seismic information is contained in a single shotpoint. Besides, the reference model is also one-dimensional : it only depends on the depth. The unknowns of the problem are the acoustic impedance disturbance and

20 velocity distribution as a function of depth. On the other hand, unlike the first application example, the seismic excitation mode (and notably the seismic wavelet) are assumed to be known. It is an inverse problem made linear by means of linearization.

Surface seismic data are often contaminated by multiple reflections which may have a rather high amplitude. Their kinematics, which is characterized by moveout in the 1D context, is generally quite different from that of primary reflexions : these two types of waves have generally travelled through geologic layers having different propagation velocities. This is the reason why linearized inversion is in itself an efficient technique for attenuating multiple reflections : the kinematics of primary reflections (the events modelled by Born's approximation) is entirely defined by the velocity distribution in the reference model : the modelled events therefore cannot adjust to multiple reflections if the moveout thereof differs from that of the primary reflections. However, because of the stationarities at the zero offset, there is in practice a partial adjustment and the linearized inversion results appear to be contaminated by the presence of the correlated noise consisting of the multiple reflections. The seismic data of Fig.12A have been modelled by taking account of the multiple reflections (there are only 3 primary reflections which reach the zero offset successively at the times : 500 ; 1000 ; 1650 ms, the 2nd arrival having a particularly high amplitude).

In the seismic response of the model obtained after conventional linearized inversion (see Fig.12B), the high-amplitude multiple arriving at approximately 1500 ms at the zero offset has been only weakly attenuated and the model is contaminated by the multiple. It is then natural to try to use the possibilities afforded by Nemeth's method in order to better discriminate multiples and primaries, notably with the improvements provided by our method.

We describe hereafter the implementation of the method for 1D linearized inversion of surface seismic data.

Experimental protocol

It consists in specifying :

- the seismic excitation mode : in our experiments, the seismic source is a source point time-modulated by a function (seismic wavelet) denoted by $w(t)$,
- 5 - the various locations where the receivers are positioned : in our experiments, the pickups are arranged at the depth 10m and cover the offset interval $[x_{\min}=0, x_{\max}=1500\text{m}]$, with a receiver every $\Delta x=100\text{m}$; we thus obtain seismic traces numbered from 0 to I,
- the pickups measure the pressure field resulting from the wave propagation,
- 10 - the recording time : : in our experiments, the pickups measure the vibrational state over the time interval $[0, T=1800\text{ms}]$, these data being sampled every $\Delta t=1 \text{ ms}$. We thus obtain samples numbered from 0 to N.

We call d_i^n the data sample recorded for the offset $x_{\min} + i \Delta x$ and at the time $n \Delta t$.

We have of course : $T = N \Delta t$ and $x_{\max} = x_{\min} + I \Delta x$.

15 *Acquisition of data in accordance with this experimental protocol*

Our experiments were carried out from synthetic data : these data were obtained by numerical resolution (finite-difference method) of the 2D wave equation :

$$\frac{1}{\sigma c} \frac{\partial^2 P}{\partial t^2} - \nabla \frac{\sigma}{c} \bar{\nabla} P = \delta(x, z - z_s) w(t) \text{ in } \mathcal{R} \times \mathcal{R}^{*+} \times [0, T]$$

with the boundary condition :

20 $P(z=0)=0$

and the initial conditions :

$$P(t = 0) = \frac{\partial P}{\partial t}(t = 0) = 0$$

an equation wherein :

- x, z, t respectively designate the lateral coordinate, the depth and the time,
- σ and c respectively represent the acoustic impedance and propagation velocity distributions (functions of the depth).

5 The wavelet selected is a conventional Ricker wavelet of central frequency 30 Hz and the impedance σ and velocity c distributions are given in Figs.13A, 13B. The seismic response of this model therefore consists of the data represented in Fig.12A. We call F_{NL} the operator which associates the seismic data with function pair $(\sigma(z), c(z))$.

Specification of the modelling operator

10 We select a reference model described by function pair $(\sigma_0(z), c_0(z))$. In the experiments presented hereafter, we have selected $c_0(z) = c(z)$ and $\sigma_0(z) = cte (=1)$. We choose as the modelling operator the Jacobian operator of operator F_{NL} at point $(\sigma_0(z), c_0(z))$. In fact, since we have selected $c_0(z) = c(z)$, the only unknown is $\delta\sigma(z) = \sigma(z) - \sigma_0(z)$ and only the corresponding component of the Jacobian is important.

15 The modelling operator, which we call $F(\delta m)$, thus is the (linear) operator which associates with function pair $\delta m = (\delta\sigma(z), \delta c(z))$ the samples $\delta y(x_i, t^n)$ which are the components of vector $F(\delta m)$ for $x_i = x_{min} + i \Delta x$ and $t^n = n \Delta t, i=0, \dots, I ; n=0, \dots, N$.

Specification of the modelling operator associated with each correlated noise

20 We specify in this paragraph the procedure that can be used for modelling each correlated noise. Each correlated noise will be characterized by correlation directions

specific thereto, which we assume to be known, at least approximately : these correlation directions are specified by means of a field of correlation vectors $\vec{c}_j(x,t)$ of components $(c_j^x(x,t), c_j^t(x,t))$ which has to be defined at any point of the domain $[x_{\min}, x_{\max}]X[0,T]$. We can, in an equivalent way, specify the correlation directions by means

5 of a pencil of correlation lines that will represent the field lines associated with $\vec{c}_j(x,t)$.

Such correlation lines can be obtained according to the technique described in patents EP-354,112 (US-4,972,383) filed by the applicant, from a pick of several phases characteristic of the noise, the information being extended to the whole domain $[x_{\min}, x_{\max}]X[0,T]$ by means of conventional interpolation and/or extrapolation procedures, as

10 described in the aforementioned patent EP-354,112 filed by the applicant. For this second application example wherein we want to eliminate the aforementioned multiple, we pick this multiple (the amplitude peak for example), which enables us to define the arrival time variations as a function of the offset and we define a correlation line continuum by simple vertical translation of the picked line (under these conditions,

15 vector $\vec{c}(x,t)$ does not depend on t). Specification of the correlation directions amounts to specifying the propagation velocity distribution of the noise $c_j(x,t)$ that is related to the correlation vector by the relation $c_j(x,t) = \frac{c_j^x(x,t)}{c_j^t(x,t)}$

Unlike the 1st application example described above, we have here noise amplitude variations along the correlation line. We assume that we can estimate these amplitude

20 variations from a measurement on the data or by means of theoretical considerations : this measurement defines a function $g(x)$.

The noise will thus be modelled by the composition of two operators :

A transport operator, as in the first application example ;

An amplitude modulation, i.e. the multiplication of the solution to the transport equation by function $g(x)$.

If we use again the scheme given for the 1st application example (and notably if we use the same hypotheses and notations), we introduce :

- a space B_j of noise-generating functions which will be the space of support functions $\beta_j(t)$ in $[t_j^{\min}, t_j^{\max}] \subset [0, T]$ and that we extend by 0 in the whole interval $[0, T]$;

noise modelling operator T_j :

$$T_j : \beta_j(t) \in B_j \rightarrow b_j(x, t) \in D$$

solution to the transport equation

$$\bar{\nabla} b_j(x, t) \cdot \vec{c}_j(x, t) = 0 \text{ in } [x_{\min}, x_{\max}] \times [0, T]$$

and meeting the initial conditions :

$$\begin{cases} b_j(x, t = 0) = 0 \\ b_j(x = x_{\min}, t) = \beta_j(t) \end{cases}$$

In fact, the transport equation is solved by means of a numerical scheme. It is for example possible to use the conventional finite-difference centered scheme. We therefore select discretization intervals Δx_j and Δt_j (which are not necessarily equal to Δx and Δt , but which we assume to be submultiples of these quantities, even if more general situations can be considered) and we introduce a grid whose nodes are the points of coordinates $(x_{\min} + i' \Delta x_j, n' \Delta t_j)$ with i' and $n' \in \mathbb{N}$. The conventional finite-difference centered scheme explained hereunder allows, from the initial conditions, to calculate step by step the different values taken by function $b_j(x, t)$ at the various nodes of the grid (to simplify the notations, we have left out subscript j associated with the correlated noise considered).

$$\frac{\xi_{i'+\frac{1}{2}}}{2\Delta t'} \left[b_{i'+1}^{n'+1} + b_{i'}^{n'+1} - b_{i'+1}^{n'} - b_{i'}^{n'} \right] + \frac{1}{2\Delta x' \tau^{n'+\frac{1}{2}}} \left[b_{i'+1}^{n'+1} + b_{i'+1}^{n'} - b_{i'}^{n'+1} - b_{i'}^{n'} \right] = 0$$

In the above formula, quantity $\alpha_{i,n'}$ represents the evaluation of quantity α (α is a generic term) at the point of coordinates $(x_{\min} + i'\Delta x_j', n'\Delta t_j')$.

Here, discretization intervals $\Delta x_j'$ and $\Delta t_j'$ have to be selected sufficiently small in relation to the wavelength because we want the numerical scheme to give a precise approximation of the function solution to the transport equation (we want no dispersion).

The last stage of the noise modelling procedure consists in selecting the samples $b_{i,n'}$ that belong to the nodes common to the two grids and in multiplying them by $g(i\Delta x)$: we thus obtain the various components of the vector (of the data space) $T_i(\beta_j)$.

Specification of a norm or of a semi-norm in the data space

A first choice would consist in taking as the norm in the data space any discrete norm meant to be an approximation (via a quadrature formula) to the norm defined on a continuous basis by :

$$\|u\|_{\mathcal{D}} = \left(\int_{x_{\min}}^{x_{\max}} dx \int_0^T dt \frac{1}{\tau(t)} u^2(x, t) \right)^{\frac{1}{2}}$$

We can also use the semi-norm, which is a better choice as we shall see in the next paragraph :

$$\|u\|_{\mathcal{D}} = \left(\Delta x \Delta t \sum_{i=0}^{I-1} \sum_{n=0}^{N-1} \frac{1}{r^{n+\frac{1}{2}}} (u_i^{n+1} + u_i^n)^2 \right)^{\frac{1}{2}}$$

where u is any vector of the data space (i and n are the indices representing the sample numbers respectively in space and in time).

Besides, other choices are also possible.

Specification, in each noise-generating functions space, of a norm for which each noise modelling operator establishes an isometric relation (or the approximation to an isometric relation) between the noise-generating functions space and the data space (provided with the semi-norm defined in §5)

We specify in this paragraph the procedure that can be used to define the norm in the noise-generating functions space associated with each correlated noise.

A first choice would consist in providing each noise-generating functions space B_j with any discrete norm meant to be an approximation (via a quadrature formula) to the norm defined on a continuous basis by (here again, we have left out subscript j to simplify the notations) :

$$\|\beta\|_B = \sqrt{\left(\int_{x_{\min}}^{x_{\max}} dx g^2(x)\right) \left(\int_0^T dt \frac{1}{\tau(t)} \beta^2(t)\right)}$$

in which case, using the fact that $b_j(x,T)=0$ for $x \in [x_{\min}, x_{\max}]$, linear operator T_j performs an isometry between B_j and D . However, since calculation of $T_j \beta_j$ is carried out numerically, the isometric character of operator T_j will not be perfect but only
 5 approximate, the approximation being all the better as discretization intervals Δx , Δt , Δx_j , Δt_j are small. The fact that we then have only approximately an isometry will lead to lower performances concerning the algorithmic implementation presented in paragraph 8.

A better choice consists in providing each noise-generating functions space B_j with
 10 the semi-norm (here again, we have left out subscript j to simplify the notations) :

$$\|\beta\|_B = \sqrt{\Delta x \Delta t \left(\sum_{i=0}^{I-1} g^2(i\Delta x)\right) \left(\sum_{n=0}^{N-1} \frac{1}{\tau^{n+\frac{1}{2}}} (\beta^{n+1} + \beta^n)^2\right)}$$

In this case, operator T_j rigorously performs an isometry between B_j and D , at least if the solution to the discretization scheme of the transport equation is identically zero for $n=N$.

15 *Definition of the cost function*

We define the cost function by the formula as follows :

$$C(\delta m, \beta_1, \beta_2, \dots, \beta_J) = \left\| F_{NL}(m_0) + F(\delta m) + \sum_{j=1}^J T_j \beta_j - d \right\|_D^2$$

It can be noted that, in the above function, $F_{NL}(m_0)$ is a constant vector.

Seeking the model and the noise-generating functions minimizing the cost function, this search being carried out by an algorithmic method taking advantage, explicitly or
5 *implicitly, of the isometry properties of the noise modelling operators (see §6)*

The algorithmic method uses the specificity of Schur's complement, a specificity associated with the isometric character of operators T_j . This can be done for example by realizing that minimizing $C(\delta m, \beta_1, \beta_2, \dots, \beta_J)$ amounts to minimizing

$$C'(\delta m) = C(\delta m, \tilde{\beta}_1(\delta m), \tilde{\beta}_2(\delta m), \dots, \tilde{\beta}_J(\delta m))$$

10 where $(\tilde{\beta}_1(\delta m), \tilde{\beta}_2(\delta m), \dots, \tilde{\beta}_J(\delta m))$ minimizes $C(\delta m, \beta_1, \beta_2, \dots, \beta_J)$ for given δm . It can be noted that the Hessian associated with this quadratic form is Schur's complement. $C'(\delta m)$ can be minimized by means of any optimization method ; cost function $C'(\delta m)$ being quadratic, the use of a conjugate gradient method is here particularly appropriate.

Determination of $(\tilde{\beta}_1(m), \tilde{\beta}_2(m), \dots, \tilde{\beta}_J(m))$ is very easy thanks to the isometric
15 character of operators T_j , notably if choices (4) and (5) have been respectively selected to define $\| \cdot \|_D$ and $\| \cdot \|_{B_j}$. If there is only one type of noise ($J=1$), which is the case for our illustration, and if the isometry is perfect (which is also the case thanks to the choices made), determination of $\tilde{\beta}_1(m)$ simply requires evaluation of the gradient of the quadratic form. If there is a superposition of several noises ($J>1$), various algorithms

can be considered to take advantage of the isometric character of operators T_j as explained for the first application example.

Figure 14B shows the results obtained (seismic response of the solution model) by implementing the method when function $g(x)$ is selected in a very simple form (affine
5 function). The improvement in relation to the conventional inversion result can be observed (Fig.14A).

Implementation examples where the physical parameter whose subsurface distribution is modelled is the acoustic impedance have been described. It is clear that the method, in its most general definition, can be applied for seeking the distribution of
10 physical quantities affected by correlated noises in any heterogeneous medium.